

UNSUPERVISED TRANSIENT LIGHT CURVE ANALYSIS VIA HIERARCHICAL BAYESIAN INFERENCE

N. E. SANDERS¹, M. BETANCOURT², AND A. M. SODERBERG¹

¹ Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA; nsanders@cfa.harvard.edu

² Department of Statistics, University of Warwick, Coventry CV4 7AL, UK

Received 2014 April 14; accepted 2014 December 12; published 2015 February 5

ABSTRACT

Historically, light curve studies of supernovae (SNe) and other transient classes have focused on individual objects with copious and high signal-to-noise observations. In the nascent era of wide field transient searches, objects with detailed observations are decreasing as a fraction of the overall known SN population, and this strategy sacrifices the majority of the information contained in the data about the underlying population of transients. A population level modeling approach, simultaneously fitting all available observations of objects in a transient sub-class of interest, fully mines the data to infer the properties of the population and avoids certain systematic biases. We present a novel hierarchical Bayesian statistical model for population level modeling of transient light curves, and discuss its implementation using an efficient Hamiltonian Monte Carlo technique. As a test case, we apply this model to the Type IIP SN sample from the Pan-STARRS1 Medium Deep Survey, consisting of 18,837 photometric observations of 76 SNe, corresponding to a joint posterior distribution with 9176 parameters under our model. Our hierarchical model fits provide improved constraints on light curve parameters relevant to the physical properties of their progenitor stars relative to modeling individual light curves alone. Moreover, we directly evaluate the probability for occurrence rates of unseen light curve characteristics from the model hyperparameters, addressing observational biases in survey methodology. We view this modeling framework as an unsupervised machine learning technique with the ability to maximize scientific returns from data to be collected by future wide field transient searches like LSST.

Key words: methods: numerical – methods: statistical – supernovae: general – surveys

1. INTRODUCTION

The majority of luminous transients in the universe are core-collapse supernovae (CC-SNe), marking the explosive deaths of massive stars (Heger et al. 2003; Smartt 2009). Stellar evolution theory, as well as both detailed observations of the explosive transient and fortuitous pre-explosion observations of the progenitor star, point to progenitor initial mass as the primary factor determining stars’ eventual death state. Metallicity, rotation rate, binarity, and other properties play important secondary roles, and permutations of these parameters are likely responsible for the extreme diversity of CC-SNe phenomenology observed in the universe (Heger et al. 2003; Smartt 2009; Smith et al. 2011; Ekström et al. 2012; Jerkstrand et al. 2014). The progenitor star mass distribution for each SN type, as well as the distribution of these secondary factors, have far reaching implications throughout astrophysics, influencing the theory of stellar evolution (Groh et al. 2013), galactic chemical evolution (Nomoto et al. 2013), hydrodynamic feedback in galaxy formation (Stilp et al. 2013), and astrobiology (Lineweaver et al. 2004).

Studies of individual transients typically focus on well observed cases within each object class, capitalizing on the availability of detailed and high signal-to-noise observations to facilitate comparison to finely tuned hydrodynamic explosion simulations and analytic light curve models (e.g., Mazzali et al. 2003; Utrobin & Chugai 2008). Syntheses of these observations, studies of large samples of SNe of a given class, are then typically composed of samples culled from these well observed cases (see, e.g., Nomoto et al. 2006; Bersten & Hamuy 2009; Jerkstrand et al. 2014). However, the properties of luminous and/or high signal-to-noise objects within a survey sample may

be systematically different from their lower luminosity/signal-to-noise counterparts, and traditional targeted transient searches themselves are inherently biased toward particular SN progenitor properties like high metallicity (Sanders et al. 2012a, 2012b). To derive truly robust and unbiased inferences about SN progenitor populations, it is therefore necessary to study transient samples in a fashion as complete and observationally agnostic as possible.

Here we discuss a methodological framework for the simultaneous modeling of multi-band, multi-object photometric observations from wide field transient surveys, which addresses certain biasing factors inherent to transient searches. This method is rooted in “hierarchical” and “multi-level” Bayesian analysis, where information about similar events within a sample is partially pooled through a hierarchical structure applied to the joint prior distribution (see Gelman et al. 2013 and references therein; see Mandel et al. 2009 for applications to SN Ia light curves). We adopt Hamiltonian Monte Carlo (HMC) as a computational technique to efficiently explore the high-dimensional and strongly correlated posterior distribution of this hierarchical model (Betancourt & Girolami 2013). The result of this modeling is simultaneous inference on physically relevant light curve parameters describing individual objects in the sample, as well as the parameter distribution among the population, regularized by the application of minimal (“weakly informative”) prior information.

In Section 2, we discuss the design and implementation of a hierarchical Bayesian model capable of simultaneously fitting large quantities of raw photometric data from wide field transient surveys to infer the population properties of the underlying SN sample. We provide a pedagogical description of hierarchical Bayesian modeling and HMC sampling, along with the complete

code specifying our transient light curve model, in the Appendix. We test this model with a sample data set of Type IIP SNe from the Pan-STARRS1 (PS1) survey (Section 3), previously published in Sanders et al. (2014). We explore the results of this test in Section 4, including comparison with inferences drawn from traditional modeling based on fits to individual light curves. We discuss the implications of this methodology for future wide field transient surveys in Section 5 and conclude in Section 6.

2. MODEL DESIGN

We have designed a hierarchical Bayesian model to simultaneously describe the individual multi-band light curves of a set of optical transients, and the population distribution of their light curve parameters. Due to the nature of the sample data set we discuss in Section 3, the nonlinear link function of our model is tailored for Type IIP SNe (see Section 2.1), but the hierarchical structure of the model is generalizable to any transient class.

Type IIP SNe are particularly apt for a hierarchical modeling approach because their long lived light curves reduce the likelihood of any individual object to have fully identifiable light curve parameters. In particular, because the plateau phase of the SN IIP light curve has a duration (~ 3 months) similar to the length of observing seasons for typical pointings of ground based telescopes, individual light curves are typically incomplete. The detected SNe IIP have often exploded between observing seasons, when their field is behind the sun, or their field sets before the plateau phase has ended. As a result, individual objects in the data set do not have the temporal coverage needed to fully identify their light curve parameters. Partial pooling among objects in the sample can compensate, helping to identify unconstrained parameter values for individual objects, while applying information from well-constrained parameters of the individual light curves to all other objects in the sample.

We note that this design is an extension of the transient light curve model we presented in Sanders et al. (2014), which was fit individually per transient to a set of photometric observations of Type IIP SNe. That model incorporated only fixed prior distributions for each parameter and did not allow for partial pooling of information between parameters of different transients. The hierarchical model presented here entails a significantly higher level of model and computational complexity than the approach of Sanders et al. (2014), and also offers distinct benefits for inference, as we will discuss below.

2.1. Light Curve Model

In Sanders et al. (2014), we have designed a physically motivated parameterized model for the SN IIP light curve, composed of 5 piecewise power law and exponential segments. The model is fully specified by a set of 12 independent parameters per optical filter and an explosion date (t_0). These parameters are four time durations (t_1, t_p, t_2, t_d) defining the knot locations of the segments, five rate parameters describing the slope of each light curve segment ($\alpha, \beta_1, \beta_2, \beta_{dN}, \beta_{dC}$), a luminosity scale (M_p), a background level (Y_b) for the photometric data, and an intrinsic scatter (V) parameter. The intrinsic scatter parameter provides flexibility in the model by admitting additional sources of measurement uncertainty to be learned from the data. The light curve model and its primary parameters are illustrated in Figure 1; a full mathematical description of the light curve model is given in Sanders et al. (2014).

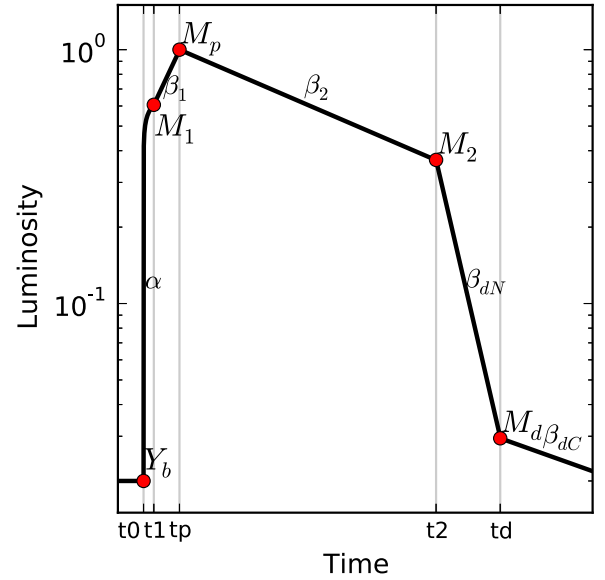


Figure 1. Schematic illustration of the five-component SN II light curve model defined in Sanders et al. (2014). The gray vertical lines denote the epochs of transition (t_x) between the piecewise components of the model. The background level (Y_b) and turnover fluxes (M_x) are marked and labeled (red points). The power law (α) and exponential (β_x) rate constant for each phase is labeled adjacent to the light curve component.

2.2. Hierarchical Structure

We allow for partial pooling between the light curve parameters of this five-component model using a linear hierarchical structure. This structure is illustrated in Figure 2. For the time (t), rate (r), luminosity scale (M), and intrinsic scatter (V) parameters of the model, this structure includes levels for individual photometric filters (e.g., r_{hF}), individual SNe-filter combinations (e.g., r_{hSNF}), and a top-level parameter (e.g., r_h); a three-level structure. For the explosion date (t_0) parameter, which is not defined per-filter, we use a modified two-level structure. The structure includes separate hyperprior distributions for objects discovered within (t_{0S}) and between (t_{0I}) observing seasons, which will have significantly different delay times between explosion and the epoch of detection. We do not adopt a hierarchical structure for the background level (Y_b) parameter, which should nominally be 0 except in the presence of artifacts among the PS1 template images used in difference imaging.

In effect, this structure means that a single top level value is drawn for each rate parameter; separate filter-level rate parameters are drawn for each of the *grizy* filters from the hyperprior distribution specified by the top level value; and bottom level rate parameters for each SN-filter combination are drawn from the hyperprior specified by the filter-level draw. In practice, this “centered” multi-level parameterization is nonoptimal because it introduces significant correlations between the hyperparameters in the model that decrease the efficiency of the MCMC sampler. Instead we use a modified, “noncentered parameterization,” where correlations between hyperparameters are exchanged for correlations between hyperparameters and data. This is a general technique applicable for any distribution in the location-scale family, and optimal when the data poorly identify the parameter values (Papaspiliopoulos et al. 2007; Betancourt & Girolami 2013). We therefore adopt normal hyperprior distributions for all the location hyperparameters in our model, and half-cauchy distributions for all scale parameters (including hyperprior width parameters; Gelman et al. 2008).

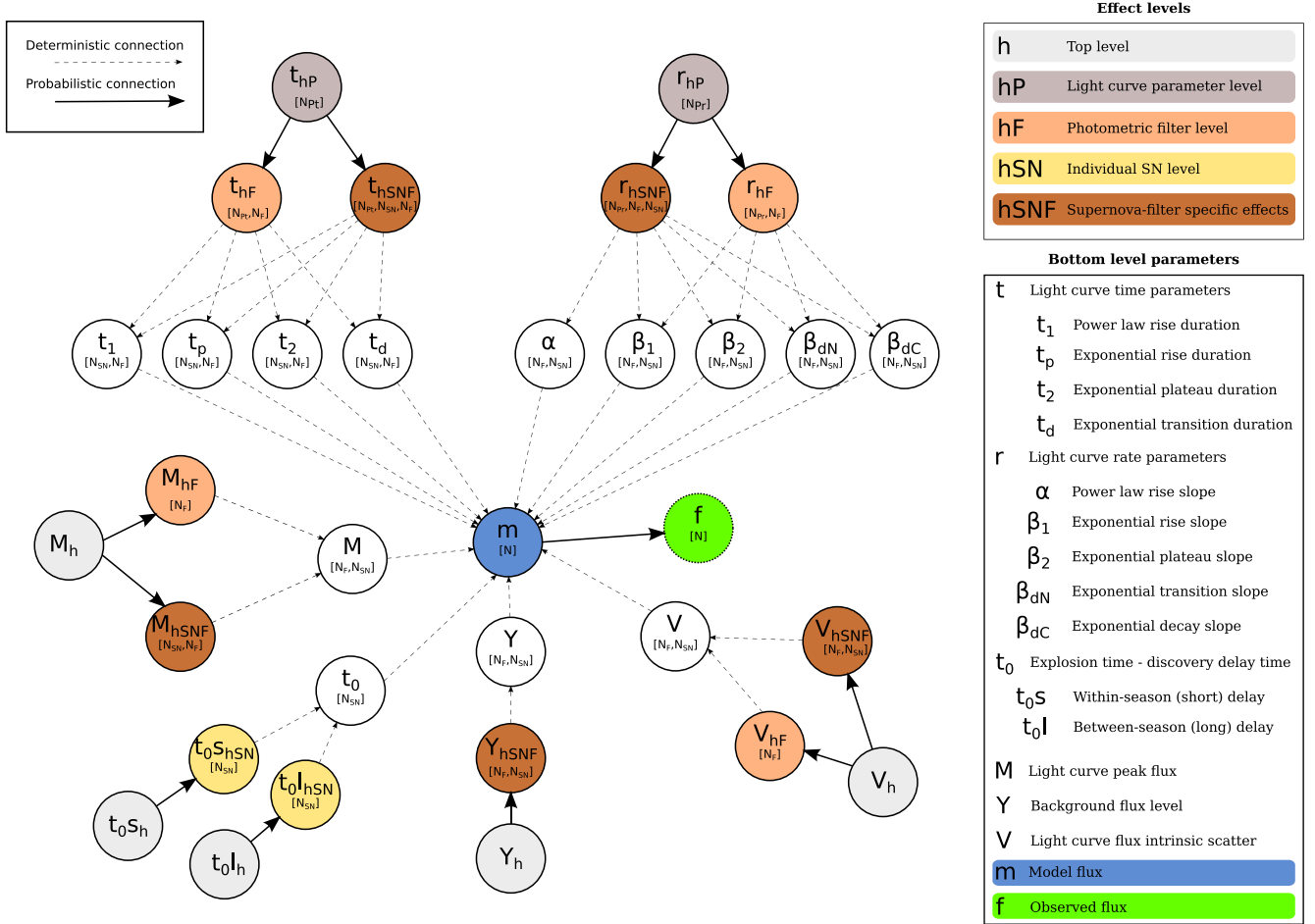


Figure 2. Directed Acyclic Graphical representation of the hierarchical structure of the multi-level light curve model parameters. The color coding of the effects levels and the meanings of the bottom level parameters are given in the key at right. The bracketed numbers indicate the dimensionality (in terms of the number of objects, N_{SN} , the number of filters, N_F , and the number of parameters in the time and rate groups, N_{Pt} and N_{Pr}) of each parameter matrix.

The hierarchical modeling framework largely eschews the specification of prior information on the parameters of the light curve model specified in Section 2.1, instead allowing the model to set its own hyperprior distributions learned from the data during fitting. We view this process, as applied to transient optical light curve studies, as a form of unsupervised machine learning. In effect, the model is learning the shape and range of variation among light curves within the transient class, given the functional form we specify for the light curve model and the Bayesian hierarchical structure, and applying that information to optimally interpret individual light curves.

However, it is necessary to set prior distributions for the top level hyperparameters, and we adopt weakly informative priors except where needed to enforce regularization of the light curve model. In particular, we assign mean values for the normal prior distribution on the filter-level parameter (t_{hF}, t_p) controlling the plateau phase rise time (t_p) to specify the within filter variation observed in Sanders et al. (2014). We do the same for the filter-level priors controlling the plateau phase rise and decay rates (β_1 and β_2). We specify the prior on the explosion date hyperparameters with means of 1 and 100 days for the within- and between-season objects, respectively. We use a restrictive cauchy (0.001) hyperprior for the top-level intrinsic scatter parameter (V_h) to regularize its ability to dominate the likelihood evaluation. We note that narrow hyperprior distributions are

needed here because the hierarchical model exponentially amplifies variances. Prior information is therefore needed to ensure a reasonable range of variation of the top level parameters and to avoid numerical overflow during sampling. The model then fits optimal values for each of these hyperparameters given the likelihood for the data, and these priors serve largely to regularize the results.

2.3. Stan Implementation

To sample from this model posterior, we employ the C++ library *Stan* (Stan Development Team 2014a), which implements the adaptive HMC *No-U-Turn Sampler* (NUTS) of Hoffman & Gelman (2014). HMC is advantageous for inference on high dimensional multi-level models because it capitalizes on the gradient of the posterior to efficiently traverse the joint posterior despite the presence of the highly correlated parameters inherent to hierarchical models (Betancourt & Girolami 2013). In practice, HMC will achieve a significantly higher effective sample size ratio (i.e., lower autocorrelation in the trace) than traditional Gibbs samplers for models with highly correlated parameters (Betancourt & Girolami 2013; Stan Development Team 2014b).

NUTS operates in two phases; “adaptation” and “sampling.” During adaptation, the algorithm automatically tunes the

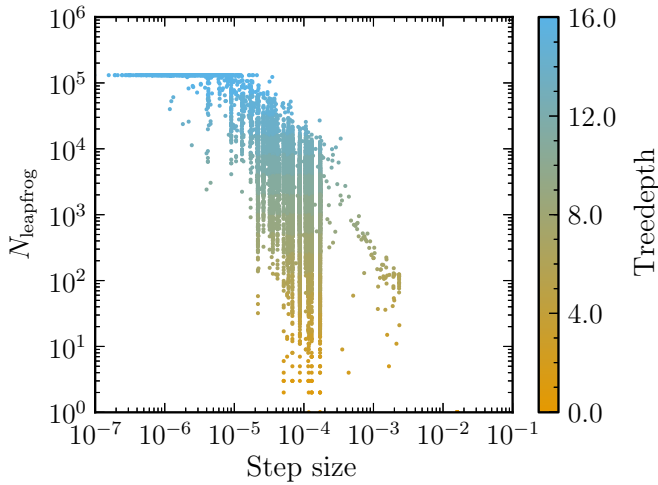


Figure 3. Illustration of the complexity of the HMC sampling procedure. The figure compares the HMC Hamiltonian discretization step size to the number of leapfrog steps (N_{leapfrog}) needed per iteration. All samples from the combined MCMC chain, including adaptation steps, are shown. The color coding shows the NUTS treedepth (key at right). Vertically correlated features indicate adapted chains (proceeding with fixed step size).

temporal step size which controls the discretization of the Hamiltonian (Hoffman & Gelman 2014). Additionally, the algorithm estimates a diagonal HMC mass matrix during adaptation, which effectively scales the global step size to the optimal value for each parameter (we do not configure *Stan* to estimate the full, “dense” mass matrix given the significant additional computational overhead). During the sampling phase, the step size and mass matrix are fixed.

We use *Stan* to construct 32 independent MCMC chains from the posterior distribution of the model.³ We have used the Harvard Faculty of Arts and Sciences “Odyssey” Research Computing cluster to run these chains in parallel, running for the cluster’s maximum job execution time of 3 days per chain, for a total utilization of 2304 cpu hrs. Given our total yield of 12,651 samples, this represents an average chain length of 395 samples and an effective sampling rate of 5.49 samples per hour per chain. For the purposes of convergence testing (Section 4.1), we consider the full chains including adaptation phase. For the purposes of light curve modeling, we exclude the adaptation phase as well as the first 20 iterations of the sampling phase, yielding 11,265 total samples from the approximate posterior stationary distribution.

The high computational cost of sampling from the model posterior distribution is due to the small HMC step size emerging from the NUTS adaptation. Figure 3 illustrates this effect, comparing the Hamiltonian discretization step size to the number of leapfrog steps per iteration as the step size varies during NUTS adaptation. As the step size decreases, the number of leapfrog steps needed (the number of posterior calculations, and therefore the execution time) grows exponentially. The horizontal feature at the top of this figure illustrates saturation of the leapfrog algorithm tree depth, suggesting that yet smaller step sizes may be needed to optimally sample from the posterior. However, given the onerous computation time required to iterate the NUTS algorithm (which is not immediately parallelizable) at the selected maximum tree depth ($\gtrsim 1$ hr of CPU time at the maximum tree depth of 16), we have elected not to increase the

maximum tree depth. As a result, the HMC sampler could potentially become stuck in local minima of the multi-dimensional posterior, biasing the resulting samples away from the tails of the true joint posterior distribution.

3. SAMPLE DATA

3.1. Pan-STARRS1 Optical Observations

Our Type IIP SN light curve sample is selected from four years of systematic Medium Deep Field observations by the PS1 telescope, as described in Sanders et al. (2014). PS1 is a high-etendue wide-field imaging system, designed for dedicated survey observations and located on a mountaintop site in the Hawaiian island chain. Observations are conducted remotely, from the University of Hawaii–Institute for Astronomy Advanced Technology Research Center (ATRC) in Pukalani. A complete description of the PS1 system, both hardware and software, is provided by Kaiser et al. (2002). The 1.8 m diameter primary mirror, 3.3 field of view, and other PS1 optical design elements are described in Hodapp et al. (2004); the array of 0.258 pixel detectors, and other attributes of the PS1 imager, is described in Tonry & Onaka (2009); and the survey design and execution strategy are described in K. C. Chambers et al. (in preparation). The PS1 Medium Deep Survey (MDS) consists of 10 pencil beam fields observed with a typical cadence of 3 days in each filter.

The PS1 observations are obtained through a set of five broadband filters, which we refer to interchangeably as g_{P1} , r_{P1} , i_{P1} , z_{P1} , and y_{P1} or simply *grizy* (Stubbs et al. 2010). MDS achieves a 5σ depth of ~ 23.3 mag in *griz* filters, and ~ 21.7 mag in the *y*-filter (with observations taken near full moon). Photometry presented here is in the “natural” PS1 system, $m = 2.5 \log(\text{flux}) + m'$, with a single zero-point adjustment m' made in each band to conform to the AB magnitude scale (Schlafly et al. 2012; Tonry et al. 2012; Magnier et al. 2013).⁴ We assume a systematic uncertainty of 1% for our PS1 observations due to the asymmetric PS1 point spread function and uncertainty in the photometric zero-point calibration (Tonry et al. 2012). The standard reduction, astrometric solution, and stacking of the nightly images is done by the PS1 IPP system (Magnier 2006; Magnier et al. 2008), and the nightly MDS stacks are processed through a frame subtraction analysis using the *photpipe* image differencing pipeline (Rest et al. 2005; Scolnic et al. 2013).

We adopt the final spectroscopic SN IIP sample from Sanders et al. (2014), including all objects sub-classified using the Support Vector Machine machine learning classification method therein. This sample consists of 18,837 total photometric data points, including 5056 robust detections, for 76 SNe IIP in the *grizy* filters. We note that the photometric observations which are not robust detections still play a significant role in the likelihood of our model, serving to constrain the rise time and decay rate parameters of the model, as well as directly identifying the background parameter Y_b . The particular transients included in this sample and their properties are described in Sanders et al. (2014).

3.2. Posterior Probability Convergence

The HMC algorithm quickly and efficiently converges on a maximal value of the global posterior probability for the model

³ The full *Stan* code for our statistical model is discussed in the Appendix.

⁴ The magnitudes quoted throughout this paper are in the AB system, except where explicitly noted.

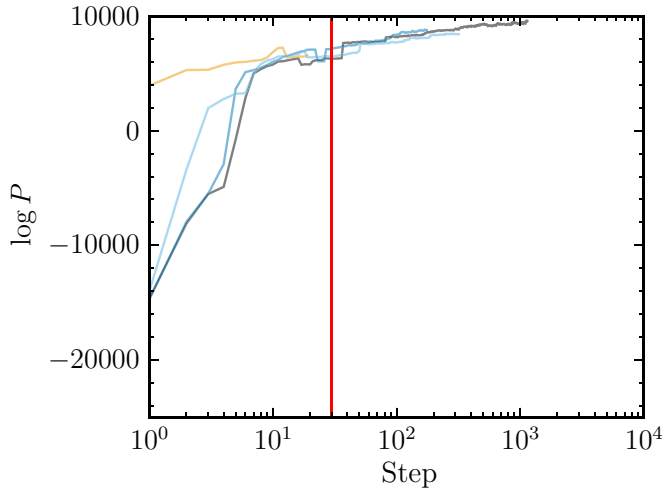


Figure 4. Markov Chain Monte Carlo trace of the global posterior probability from the hierarchical model fit, showing the fast convergence of the HMC algorithm to near the optimal parameter values. A random subset of chains from the hierarchical model fit are shown (different lines). The red vertical line marks the end of the NUTS adaptation phase, at which point the HMC step size is fixed. The probability shown on the y axis is not normalized and therefore has arbitrary units.

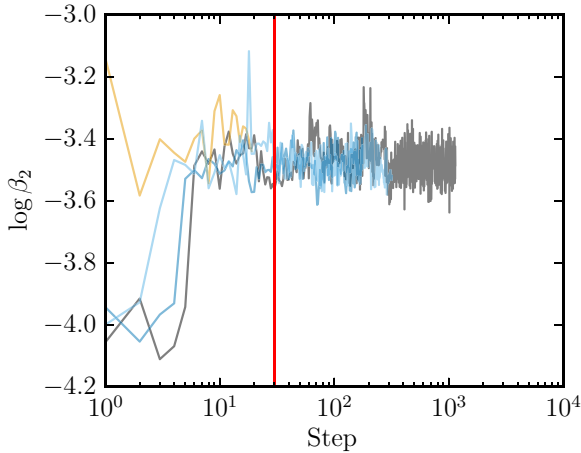


Figure 5. Illustration of convergent HMC sampling for the well identified bottom-level model parameters. Like Figure 4, but for the bottom level parameter β_2 for a randomly selected SN (r band).

by identifying optimal values for each bottom level light curve parameter for all the SNe and for the hyperparameters. Given that the global model for the PS1 SNe IIP sample has a total of 9176 individual parameters, this fast convergence is a significant testament to the efficiency of HMC as an optimization engine for high-dimensional functions.

Figure 4 shows the posterior probability evolution of the Markov chains as the NUTS sampler adapts and then reaches the sampling phase. Chains typically converge near the maximum achievable posterior probability during our warmup period of only 30 iterations.

4. RESULTS

4.1. Sampling Characteristics and Fit Convergence

Figure 5 shows the MCMC trace for a well identified bottom level parameter; the values drawn from the HMC algorithm for the plateau phase decline rate (β_2) of an object (PS1-12cey)

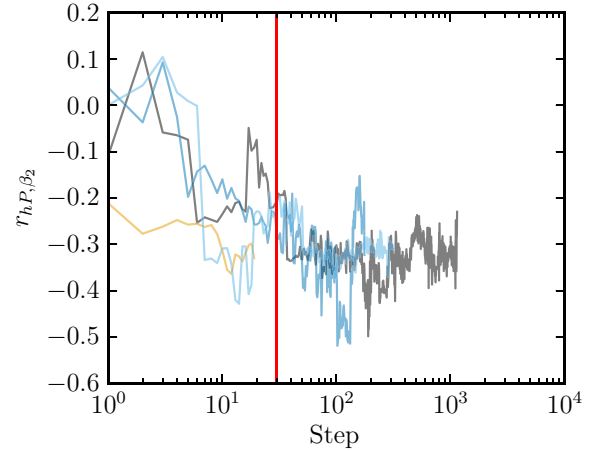


Figure 6. Illustration of slow moving HMC sampling for the higher level model parameters. Like Figure 4, but for the hyperparameter r_{hP, β_2} , the top level parameter controlling the β_2 decay rates.

with sufficient r -band photometry to constrain this phase of the light curve. The sampler moves quickly in this dimension, with low autocorrelation between samples, and the parameter is acceptably convergent (with potential scale reduction factor $\hat{R} = 1.14$).

In contrast, Figure 6 shows the trace for a moderately well identified top level parameter (r_{hP}), controlling the global plateau phase decay rate across all filters. The trace indicates that the sampler is moving more slowly in this dimension, with significant autocorrelation between samples. This top level parameter has apparently not yet converged ($\hat{R} = 1.77$). This suggests that additional sampling is needed to achieve a desirable level of convergence among some hyperparameters, but the computational cost is prohibitive at this time; we discuss alternative methods for achieving convergence with HMC in Section 5.

The origin of the convergence challenges facing the HMC algorithm are illustrated in Figure 7. The figure, showing a slice from the joint posterior distribution, illustrates the high correlation between the hierarchically linked parameters in the model. In contrast, Figure 8 shows a slice of the joint posterior along the dimensions of the top level and filter-level hyperparameters for the plateau phase time duration. Dependence between these hyperparameters was obviated via selection of the noncentered parameterization (Section 2.2) and, indeed, their marginal posteriors have very low correlation.

4.2. Posterior Predictive Check Comparison

We validate the success of our model in describing the light curve behavior of objects in our SN IIP sample using posterior predictive checks (Gelman et al. 2013), comparing the distribution of luminosities predicted under our fitted light curve model to the observed photometric data. Figure 9 shows a posterior predictive check for PS1-10zu, whose poor temporal coverage illustrates the strengths of the hierarchical model. The figure compares the r -band light curve fit for this object to the fit under the individual-level model presented in Sanders et al. (2014), which uses an identical five-component light curve model, but does not make use of partial pooling between SNe. The hierarchical fit achieves significantly greater constraints on the parameters describing the rising phases of the SN, resulting in a much tighter distribution of explosion dates and

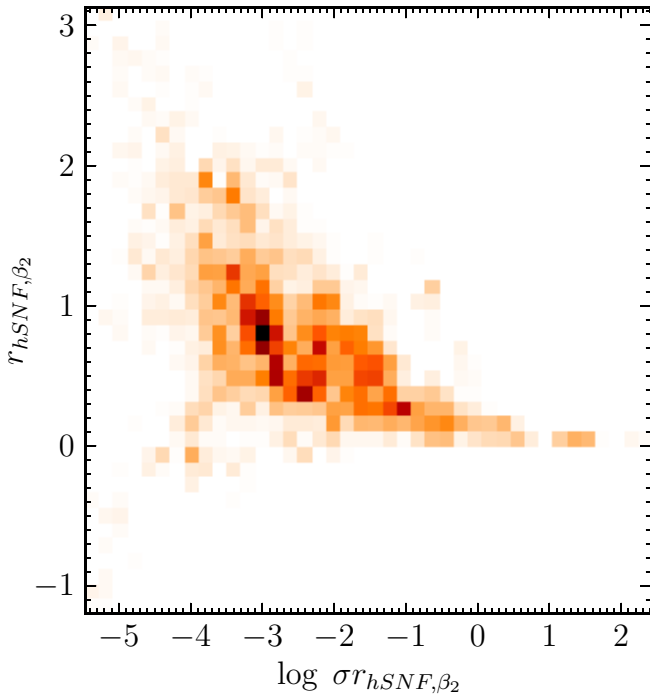


Figure 7. Illustration of the high curvature in the hierarchical model posterior. A histogram of samples from a slice of the joint posterior distribution is shown, along the dimensions of the SN-filter level mean and width (on a log scale) hyperparameters for the plateau phase decay rate (r_{hSNF, β_2}).

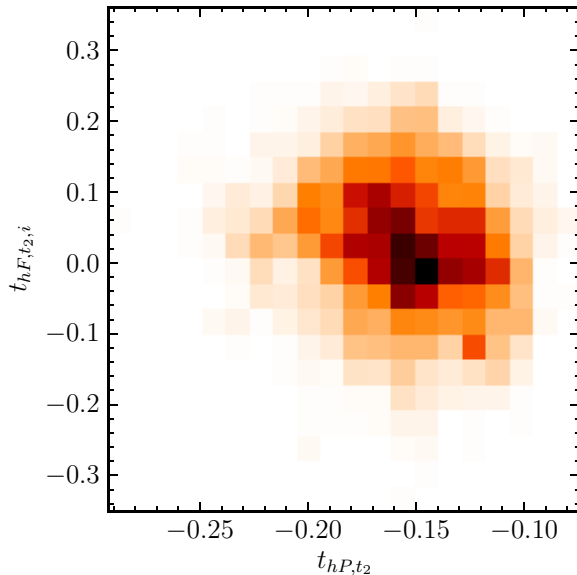


Figure 8. Illustration of the effectiveness of the noncentered parameterization in removing dependency between hyperparameters. Like Figure 7, but showing a slice along the dimensions of the top level and filter-level hyperparameters for the plateau phase time duration (t_{hP, t_2} and $t_{hF, t_2, i}$).

plateau durations (a parameter critical for physical inference on the progenitor star). The improvement is due to the strongly identified plateau duration hyperparameters (Section 4.3). In the individual-level model, a weakly informative prior distribution was established for this parameter based on the theoretically predicted range of plateau duration variation; in the multi-level model, the hyperparameters are inferred from the data themselves, resulting in much stronger prior information at the individual SN level. That the fit to the later phases of

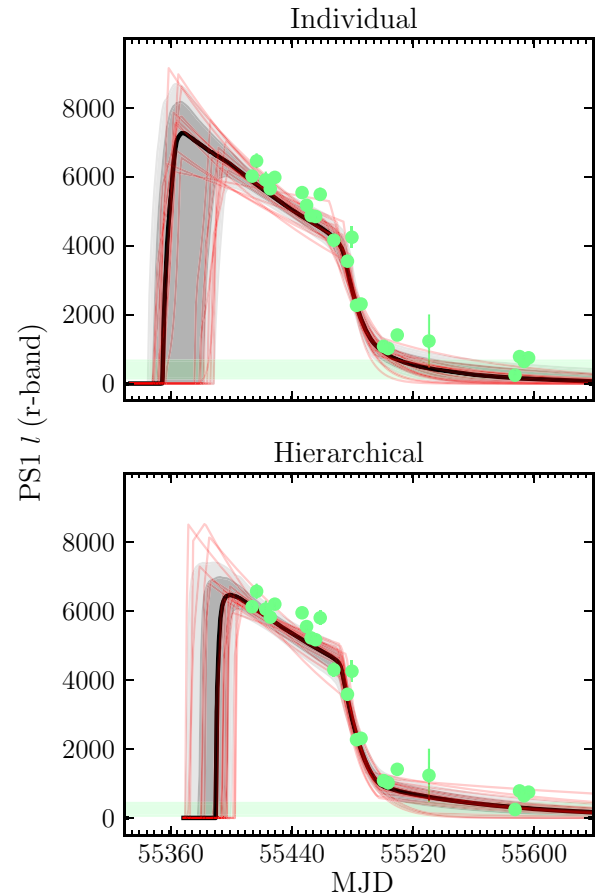


Figure 9. Comparison of posterior predictive checks between the individual light curve fit of Sanders et al. (2014), above, and the hierarchical model fit, below, for the SN PS1-10zu. Each plot shows the r -band photometry for the SN (green points and errorbars), which has detection only of the final phases of the light curve because the object exploded between observing seasons. The units of luminosity, l , are described in (Sanders et al. 2014). The shaded areas show the 1σ and 2σ confidence intervals for the posterior predictive distribution of the five-component light curve model (see Section 4.2), and the solid black line shows the median of that confidence interval. The red lines show a random subset of sampled light curve models corresponding to parameter values from the MCMC chain. The horizontal green bar shows the range fitted for the zero-point luminosity offset.

the light curve, where the data are strongly identifying, is indistinguishable from the fit obtained in the individual model is validation of the unbiased performance of the hierarchical model.

Figure 10 shows a comparison of fits for several additional objects, illustrating features of the hierarchical modeling framework under different data scenarios, which we describe here. PS1-11azd was observed only during the rise and initial stages of the plateau phase. The posterior predictive luminosity distribution of the individual and hierarchical models are similar, but the plateau decline phase duration (t_2) parameter is much more constrained in the hierarchical model (Figure 10(a)).

PS1-12bku (Figure 10(b)) was observed from explosion through the final, radioactive decay-dominated phase. Generally, this case confirms that where the data is strongly identifying, the hierarchical model produces fits in agreement with the individual-level model. Interestingly, for this object there is a r -band photometric observation with relatively high uncertainty at $\sim +80$ days, which introduces a degeneracy in the posterior—whether this point should be assigned to the plateau or transition

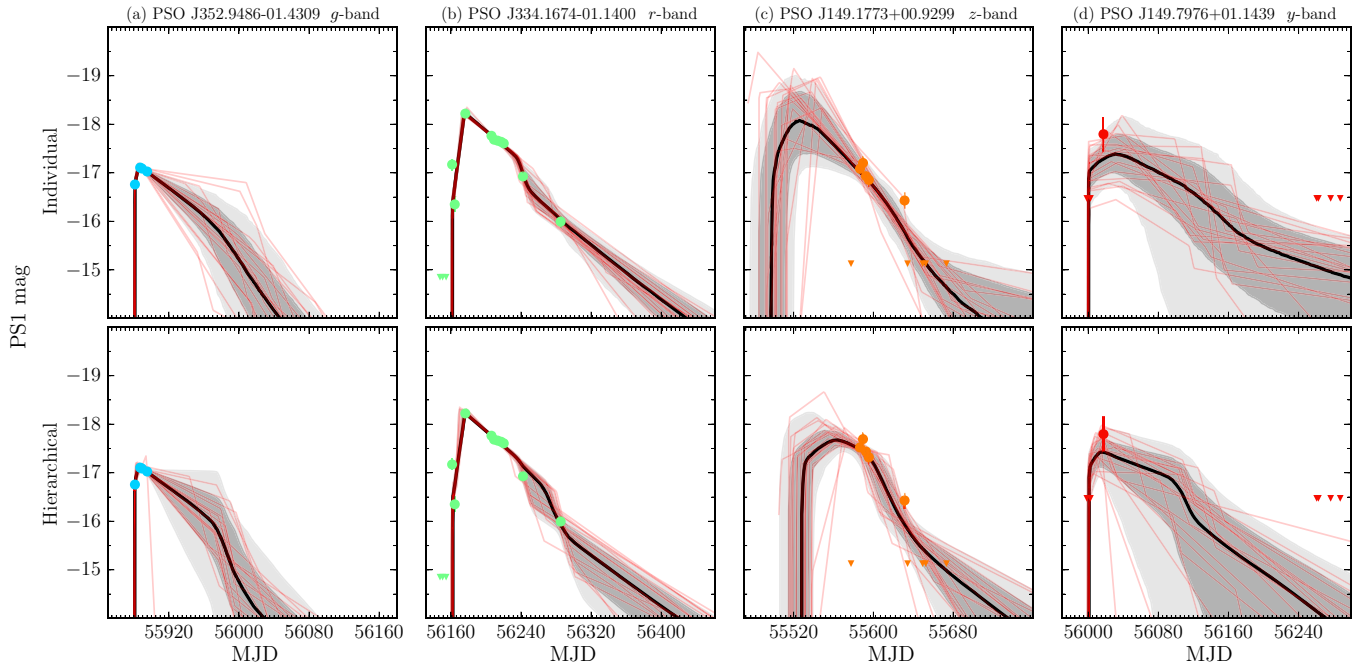


Figure 10. Illustration of the advantages provided by the hierarchical model for fitting under different data coverage scenarios, relative to individual light curve fitting. Like Figure 9, but on the absolute magnitude scale and for a variety of PS1 SNe and photometric filters. The individual model fits are shown above, the with the corresponding hierarchical fits below. Triangles represent the locations of photometry that are not robust detections; they are plotted at the y axis position of the limiting PS1 magnitude of the filter, though the actual flux and uncertainty of these points is used in the Bayesian likelihood calculation.

phase of the light curve. The fit for this SN under the hierarchical model looks similar to the individual-level model fit, exploring both forks of the degeneracy. However, the two fits favor opposite sides of the fork. The individual fit maximizes the likelihood of the r -band photometry for the object alone, placing the point on the transition phase, while the hierarchical fit prefers the solution where the point falls on the plateau. Because the fork favored by the hierarchical model is more consistent with the modeled distribution of plateau durations among SNe IIP (based on partial pooling from the other objects in the sample), it is the more well justified solution.

In Figure 10(c), the hierarchical z -band fit for the SN PS1-11ai is much more highly regularized to match the shape of other z -band light curves than the individual fit. This leads to significantly improved constraints on the peak magnitude and plateau duration for this object. For PS1-12wn (y band; Figure 10(d)), only 1 robust photometric detection is available, and all the photometry is highly uncertain ($\delta m \sim 0.5$ mag at 1σ). The individual light curve fit in this case is very poorly regularized, obeying the peak magnitude suggested by the detection and the limits suggested by the nondetections, but otherwise has very poorly constrained light curve properties like plateau duration and decline rate. The hierarchical fit in this case is far superior in regularization, showing a characteristic y -band SN IIP light curve shape matched to the available photometry.

4.3. Population Parameter Distribution Characteristics

To further the investigation of SN IIP plateau durations from Sanders et al. (2014), in Figure 11 we compare the plateau duration inferences from that project to the duration distributions inferred from the hierarchical model. The plateau duration distribution hyperprior in the hierarchical model is a sum of the t_p and t_2 hyperpriors. The hyperpriors include both location (e.g., t_{hp}) and width (e.g., $\sigma_{t_{hp}}$) hyperparameters, so we visualize the posterior distribution of hyperpriors by showing multiple

lognormal hyperpriors corresponding to random draws of the hyperparameters. We focus on the r -band durations here, and so include both the t_{hp} and t_{hF} hyperparameters.

The distribution of bottom-level plateau duration parameters for our hierarchical model agree well with the individual light curve fits from Sanders et al. (2014). Taking the median bottom level parameter from the MCMC chain, the distribution of values from the hierarchical model has a mean and standard deviation of 90 ± 6 days, compared to 92 ± 14 days for the individual model. Note that the variance in the hierarchical model distribution is significantly lower than from the individual fits because partial pooling between objects constrains the bottom level posteriors.

By directly modeling the underlying population of transients, the hierarchical modeling framework allows us to overcome potential biases in transient search methodology. In particular, although long duration SNe IIP are less likely to be observed with full temporal coverage in ground based transient searches (Section 2), we can estimate the fraction of unseen, long-duration transients in the population from the hierarchical model posterior. This product of the posterior distribution of the hierarchical model is constrained by both the observed characteristics of objects in the sample, and characteristics allowed by pre-explosion and late time nondetections from the transient search at the same location.

Our hyperparameter posterior distributions suggest there is a 60% probability that at least 10% of the underlying population of SNe IIP have r -band plateau durations longer than the bottom level parameter value for any individual object in the sample (> 112 days). The probability that at least 20% of objects fall above this value is 30%. Among the sampled hyperprior distributions, the median of the population standard deviation is 33days. The standard deviation distribution has a strong tail at larger values, shown in Figure 11. These results emphasize and support the finding of Sanders et al. (2014), that the plateau duration distribution of SNe IIP has significant variance.

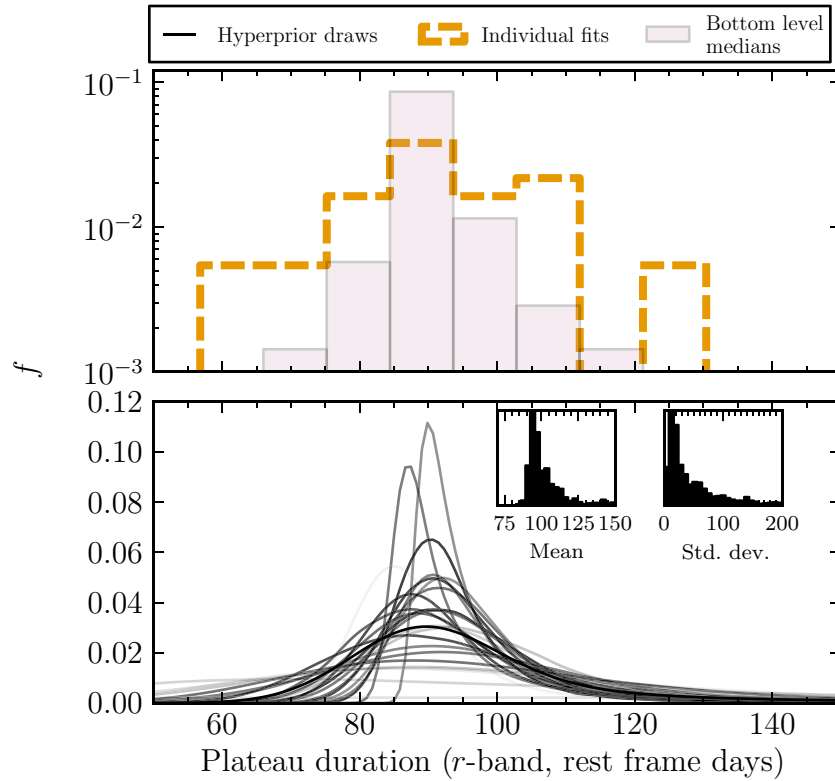


Figure 11. Comparison of population level inferences on the SN IIP plateau duration distributions (in r band) between the hierarchical and individual model fits. Top: the dashed line shows the histogram of plateau durations measured in Sanders et al. (2014) using the individual model (including only the objects with well constrained posteriors, as defined therein). The shaded area shows a histogram of the bottom level plateau duration parameters under the hierarchical model (for all objects in the sample). Bottom: the solid lines show draws of the plateau distribution hyperprior (a sum over the t_p and t_2 hyperpriors) from the hierarchical model fit, with the opacity scaled to the probability of the draw under the posterior. The inset plots show the marginalized posterior distributions for the mean and standard deviation of the summed hyperprior distributions (in rest frame days).

While we have shown that the hierarchical modeling methodology can retrieve information about individuals in the SN population even when their light curves are not fully observed, it is still subject to bias due to the systematic absence of individuals from the sample. In particular, transients with short durations and/or dim absolute magnitudes are less likely to be detected in ground based transient searches. In the present SN IIP sample, this bias is somewhat moderated by the result of Sanders et al. (2014) that brighter objects of this class also have faster light curve decline rates; therefore, short-duration SNe IIP exploding after the end of an observing season are more likely to still be visible in their post-plateau decay phase during the next observing season. This bias could be assessed and quantified using simulations on a synthetic light curve population with known duration and peak magnitude distribution, and could potentially be modeled as a truncation mechanism on the light curve parameter hyperpriors to alleviate this bias in the modeling results.

5. DISCUSSION

The multi-level model developed and applied in this paper points to a methodological framework for the interpretation of SN light curves from the next generation of wide-field transient searches, such as the Large Synoptic Survey Telescope (LSST; LSST Science Collaboration et al. 2009). In the coming era, the volume of available photometric data will increase dramatically, while the human and observational resources for follow-up of individual objects will not. In this regime, to capitalize

on the larger SN sample sizes afforded by these next generation searches, it will be critical to apply population level light curve modeling. To do this, it is necessary to apply analytical methods that are robust to data sparseness and posterior nonidentifiability for individual objects, and computational methods that are capable of generating inferences from large data sets. The combination of hierarchical Bayesian methodology and HMC methods explored in this work are natural methods for addressing both these concerns.

This work also suggests future paths for improvement of the Bayesian light curve modeling framework presented here.

First, to permit applications to purely photometric data sets (where the SN classification and redshift are not known or poorly constrained), the model must be generalized. The redshift can trivially be added to the multi-level model as a vector of free parameters, but it will introduce significant posterior correlations and interactions that will need to be tested and may require the assertion of significant prior information to aid identifiability.

Second, Additional model components are needed to allow for application to multiple SN classes simultaneously. This could be accomplished through categorical mixture modeling with several different, physically motivated SN light curve prescriptions for each SN type or by using a more generic, nonparametric, or continuously expanding light curve model to allow fitting of diverse SN types. Both of these solutions would support classification inferences, by assessment of the categorical simplex parameter posteriors directly, or by a clustering analysis of the continuous model expansion parameters, respectively.

Third, applications to transient searches delving to significantly higher redshifts (and also the nonparametric modeling approach) may require full three-dimensional modeling of the SN spectral energy distribution evolution, rather than two-dimensional light curve modeling in each filter, in order to permit K -corrections at all distances.

Fourth, in order to capture known dependencies between light curve parameters (for example, the peak magnitude—plateau-phase decline rate relation presented in Sanders et al. 2014 and Anderson et al. 2014), the full covariance structure of the light curve parameters should be modeled. Such modeling can also reveal unanticipated correlations between light curve parameters. For example, Mandel et al. (2011) have explicitly modeled the covariance between the peak magnitude and decline rate parameters of Type Ia SNe.

Fifth, incorporating host galaxy information will be critical to producing purely photometric informative inferences across SN classes. This should include modeling of the distribution of host galaxy global properties per SN class to uncover and take advantage of differences in progenitor properties (see, e.g., Kelly & Kirshner 2012; Sanders et al. 2012a; Foley & Mandel 2013; Lunnan et al. 2014; McCrum et al. 2014), as well as the line-of-sight host galaxy extinction and reddening law (e.g., Mandel et al. 2011).

Sixth, we note that replicated data simulation is a useful framework for model checking of generative Bayesian models like that presented here and the extensions discussed above. With fixed values chosen for the hyperparameters of the model, the *Stan* code we present in the Appendix can be used to generate replicated SN light curve data from the hyperprior distributions. With hyperparameters matched to the mode of the inferred posterior of our PS1 SN IIP test data set, the properties of the replicated data can then be compared to the original data set. This would demonstrate, for example, that the plateau phase decay rate and peak luminosity relation among SNe IIP is not reflected in the model structure (as discussed above). With hyperparameters chosen to more tightly constrain the light curve parameter prior distributions, more subtle assumptions of the model can be examined. In particular, we anticipate that comparison of such replicated data to highly constraining nearby SN IIP light curves will reveal parameters for which normal prior distributions are poorly suited, being skewed or otherwise nonGaussian (see, e.g., Mandel et al. 2014). Such simulations, providing highly identifying replicated data, would also be useful for testing the convergence properties of the model and HMC sampler under conditions like will be produced by upcoming wide field survey programs.

Finally, we look to Riemannian manifold Hamiltonian Monte Carlo (RMHMC; Betancourt 2013) techniques for permitting posterior characterization in the future, in the face of these additional modeling complexities. Compared to traditional, Euclidian Hamiltonian Monte Carlo (EHMC; as applied for this work), RHMC samplers efficiently explore highly correlated and high-dimensional posterior functions by automatically adapting the Hamiltonian integration step size to a value optimal for local conditions (Girolami & Calderhead 2011; Betancourt & Girolami 2013). This capability would permit unbiased sampling even amidst models with joint posterior distributions with higher curvature than the multi-level model examined here, such as a model including interactions between filter-level parameters or interactions between light curve parameter groups (e.g., t - r interactions). Our attempts to fit such a model with EHMC have not achieved convergence within reasonable

integration times, with the high posterior curvature preventing the hyperparameters from moving at a sufficient rate to produce convergent chains. The addition of RHMC sampling capabilities to *Stan* in the near future (Betancourt 2013) will make these techniques accessible to astronomers in the LSST era.

6. CONCLUSIONS

We have explored the use of Bayesian hierarchical modeling and HMC to enable population-level inference on multi-band transient light curves from comprehensive analysis of optical photometry from wide field transient searches. The primary conclusions of this work are as follow.

1. While computational limits still challenge the implementation of hierarchical models, due to the high curvature in their joint posterior distributions, sufficient convergence is achieved in the bottom level model parameters (Section 4.1) to enable their immediate application for transient light curve studies.
2. Comparisons between light curve posterior predictive distributions from our hierarchical model fit to the individual light curve fits of Sanders et al. (2014) show strong agreement for well identified parameters, and show an advantage for hierarchical models among poorly identified parameters (Section 4.2). In particular, partial pooling of parameter information between transients supports improved regularization of light curve shapes, and supports model selection between partially degenerate light curve parameter scenarios.
3. By directly modeling the underlying transient population, hierarchical models permit inference on the occurrence of properties not observed within the data set (Section 4.3). This feature is of particular value in overcoming observational biases induced by ground based transient searches, such as the under-representation of long duration transients like some SNe IIP.

We have concluded with a discussion of future directions for this modeling (Section 5), including applications to upcoming wide field transient searches, extensions to the hierarchical model structure developed here, and expanded capabilities to be enabled by the advent of RMHMC.

We thank K. Mandel for sage guidance and many helpful conversations; M. Brubaker, B. Carpenter, A. Gelman, and the *Stan* team for their excellent modeling language and HMC sampler and for thoughtful feedback on our model design; and the entire PS1 collaboration for their monumental efforts toward the collection of the SN IIP light curve data set.

The PS1 surveys have been made possible through contributions of the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation under grant No.

AST-1238877, the University of Maryland, and Eotvos Lorand University (ELTE). Support for this work was provided by the David and Lucile Packard Foundation Fellowship for Science and Engineering awarded to A.M.S. M.B. is supported under EPSRC grant EP/J016934/1. Computations presented in this paper were performed using the Odyssey supercomputing cluster supported by the FAS Science Division Research Computing Group at Harvard University.

Facility: PS1

APPENDIX

HIERARCHICAL LIGHT CURVE STAN MODEL

Below we reproduce the full hierarchical model for the five component piecewise SN IIP light curves in the *Stan* modeling language, as described in Section 2. The *Stan* model specification format is documented in the *Stan* Modeling Language Users Guide and Reference Manual (Stan Development Team 2014b).

First, we provide a pedagogical review and description of the hierarchical Bayesian modeling framework and the HMC sampling algorithm. For a more complete introduction, see Gelman et al. (2013), Betancourt & Girolami (2013), and Stan Development Team (2014b).

In Bayesian models, the posterior probability of a parameter value, θ , is modeled as proportional to the product of the likelihood of the observed data, y , given the parameter value and the prior probability of the parameter value:

$$\pi(\theta|y) \propto \pi(y|\theta)\pi(\theta). \quad (\text{A1})$$

In a hierarchical Bayesian model, the parameters of the prior distribution are themselves drawn from a “hyperprior” distribution, requiring an additional product term in the posterior probability and additional “hyperparameters,” τ :

$$\pi(\theta, \tau|y) \propto \pi(y|\theta, \tau)\pi(\theta|\tau)\pi(\tau). \quad (\text{A2})$$

Additional hyperprior distributions can be incorporated to deepen the hierarchy for the parameter θ or to establish hierarchies for different parameters.

For example, in our transient light curve model, the rise rate parameters for each SN i and filter j , $\alpha_{i,j}$, are drawn from a filter-level prior distribution (controlled by the parameter α_{hF}), which is in turn drawn from a top-level hyperprior distribution (controlled by the parameter α_h).⁵ We model each prior level with a normal distribution (N) having standard deviations (σ_x) drawn from half-Cauchy distributions (C), as follows:

$$\begin{aligned} \alpha_{i,j} &\sim N(\alpha_{hF,j}, \sigma_{\alpha_{hF,j}}) \\ \alpha_{hF,j} &\sim N(\alpha_h, \sigma_{\alpha_h}) \sigma_{\alpha_{hF,j}} \sim C(0, \sigma_{\alpha_h}) \\ \alpha_h &\sim N(0, 1) \quad \sigma_{\alpha_h} \sim C(0, 0.1). \end{aligned} \quad (\text{A3})$$

We establish similar hierarchies for the other parameters in the light curve model, as illustrated in Figure 2.

The HMC algorithm uses information from the gradient of the posterior distribution to improve the efficiency of Markov Chain Monte Carlo simulations. HMC proceeds by simultaneously sampling parameter vectors q and a vector of auxiliary

parameters p during the Markov Chain simulations. These vectors can be thought of equivalent to the position and momentum vectors in Hamilton’s dynamics equations, familiar from classical mechanics:

$$\begin{aligned} \frac{dq}{dt} &= \frac{\partial T}{\partial p} T(p|q) = -\log \pi(p|q) \\ \frac{dp}{dt} &= -\frac{\partial T}{\partial q} - \frac{\partial V}{\partial q} \quad V(q) = -\log \pi(q), \end{aligned} \quad (\text{A4})$$

where the change in time, dt , is manifested as the iterations of the Markov Chain simulation. Effectively, HMC simulates the movement of a particle through a high-dimensional space with potential energy defined by the posterior distribution of the model and kinetic energy drawn randomly from a pre-defined momentum distribution. The distribution of the vector q over the course of the simulation then provides parameter samples from the posterior distribution.

In common applications of HMC, a normal distribution is employed for the momentum, with covariance independent of the position vector. This variant is known as Euclidean HMC. The convergence properties of the simulation are sensitive to the covariance “tuning” parameters chosen for the momentum distribution, which effectively sets the position-space step size scale of the simulation. The *No-U-Turn Sampler* algorithm of Hoffman & Gelman (2014) provides an adaptive mechanism for optimally selecting these tuning parameters during a “warmup” run of the HMC simulation.

Our hierarchical transient light curve model takes the following data as input: `N_obs`, the total number of photometric data points; `N_filt`, the number of photometric filters; `t`, a vector of MJD dates of the photometric observations; `fL`, a vector of luminosities corresponding to the photometric observations (with units as described in Sanders et al. 2014); `dfL` a corresponding vector of luminosity uncertainties; `z` the redshift; `t0_mean` an initial estimate of the explosion date (for initialization and for centering the explosion date prior distribution); `J` a vector of integers specifying the filter ID of each photometric observation; `Kcor_N`, a matrix of pre-computed K -corrections for each filter, in magnitudes with spacing of 1 day; `fluxscale` the zero-point of the luminosity unit system (`fluxscale` = 10^7 in the system we have employed); and `duringseason`, a boolean value specifying whether the object exploded within or between observing seasons, for selection of the explosion date prior distribution parameters. The calculation of the model light curve flux and application of the K -correction values is performed in the `transformed parameters` section, and the prior and likelihood calculations are performed in the `model` section. Certain vector-valued prior distribution parameters are specified in the `transformed data` section for convenience. We note that the higher level parameters for the five different light curve rates and four different phase durations are grouped together in vectors (e.g., r_{hP} and t_{hP} for the top level, and r_{hF} and t_{hF} for the filter level, respectively) for convenience.

The *Stan* model is then compiled and run (Stan Development Team 2014b) to yield MCMC samples from the posterior distribution of light curve parameters. We configured the *No-U-Turn Sampler* to use fixed 0 initialization of the parameter values, an adaptation phase of 30 steps, a maximum treedepth of 16, and otherwise employed the default sampler parameters. We have used CmdStan version 2.2.0.⁶

⁵ As discussed in Section 2.2, our *Stan* model actually employs a noncentered parameterization which draws the filter and SN-filter parameters independently of each other. This aids computational efficiency, but the two formulations are probabilistically equivalent.

⁶ <https://github.com/stan-dev/stan/releases/tag/v2.2.0>

```

data {
  int<lower=0> N_obs;
  int<lower=0> N_SN;
  int<lower=0> N_filt;
  vector[N_obs] t;
  vector[N_obs] fL;
  vector[N_obs] dfL;
  vector[N_SN] z;
  vector[N_SN] t0_mean;
  int<lower=1,upper=N_filt> J[N_obs];
  int<lower=1,upper=N_SN> SNid[N_obs];
  int<lower=0> Kcor_N;
  real Kcor[N_SN, N_filt, Kcor_N];
  real<lower=0> fluxscale;
  vector<lower=0,upper=1>[N_SN] duringseason;
}
transformed data {
  vector[N_filt] prior_t_hF[4];
  vector[N_filt] prior_t_hF_s[4];
  vector[N_filt] prior_r_hF[5];
  vector[N_filt] prior_r_hF_s[5];
  for (i in 1:N_filt) {
    prior_t_hF[1,i] <- 0;
    prior_t_hF_s[1,i] <- 0.1;
  }
  prior_t_hF[2,1] <- -1;
  prior_t_hF[2,2] <- -0.5;
  prior_t_hF[2,3] <- 0;
  prior_t_hF[2,4] <- 0.5;
  prior_t_hF[2,5] <- 1;
  for (i in 1:N_filt) {prior_t_hF_s[2,i] <- 0.1;}
  for (i in 1:N_filt) {
    prior_t_hF[3,i] <- 0;
    prior_t_hF_s[3,i] <- 0.1;
  }
  for (i in 1:N_filt) {
    prior_t_hF[4,i] <- 0;
    prior_t_hF_s[4,i] <- 0.1;
  }
  for (i in 1:N_filt) {
    prior_r_hF[1,i] <- 0;
    prior_r_hF_s[1,i] <- 0.1;
  }
  prior_r_hF[2,1] <- 2;
  prior_r_hF[2,2] <- 1;
  prior_r_hF[2,3] <- 0;
  prior_r_hF[2,4] <- -0.5;
  prior_r_hF[2,5] <- -1;
  for (i in 1:N_filt) {prior_r_hF_s[2,i] <- 0.1;}
  prior_r_hF[3,1] <- 1;
  prior_r_hF[3,2] <- 0.3;
  prior_r_hF[3,3] <- 0;
  prior_r_hF[3,4] <- -1;
  prior_r_hF[3,5] <- -1;
  for (i in 1:N_filt) {prior_r_hF_s[3,i] <- 0.1;}
  for (i in 1:N_filt) {
    prior_r_hF[4,i] <- 0;
    prior_r_hF_s[4,i] <- 0.1;
  }
  for (i in 1:N_filt) {
    prior_r_hF[5,i] <- 0;
    prior_r_hF_s[5,i] <- 0.1;
  }
}

```

```

}
parameters {
  vector[4] t_hP;
  vector<lower=0>[4] sig_t_hP;
  vector[N_filt] t_hF[4];
  vector<lower=0>[N_filt] sig_t_hF[4];
  vector[N_SN * N_filt] t_hSNF[4];
  vector<lower=0>[N_SN * N_filt] sig_t_hSNF[4];
  vector[5] r_hP;
  vector<lower=0>[5] sig_r_hP;
  vector[N_filt] r_hF[5];
  vector<lower=0>[5] sig_r_hF[5];
  vector[N_SN * N_filt] r_hSNF[5];
  vector<lower=0>[N_SN * N_filt] sig_r_hSNF[5];
  real M_h;
  real<lower=0> sig_M_h;
  vector[N_filt] M_hF;
  vector<lower=0>[N_filt] sig_M_hF;
  vector[N_SN * N_filt] M_hSNF;
  vector<lower=0>[N_SN * N_filt] sig_M_hSNF;
  real Y_h;
  real<lower=0> sig_Y_h;
  vector[N_SN * N_filt] Y_hSNF;
  vector<lower=0>[N_SN * N_filt] sig_Y_hSNF;
  real t0s_h;
  real<lower=0> sig_t0s_h;
  vector[N_SN] t0s_hSN;
  vector<lower=0>[N_SN] sig_t0s_hSN;
  real t0l_h;
  real<lower=0> sig_t0l_h;
  vector[N_SN] t0l_hSN;
  vector<lower=0>[N_SN] sig_t0l_hSN;
  real<lower=0> V_h;
  vector<lower=0>[N_filt] V_hF;
  vector<lower=0>[N_SN * N_filt] V_hSNF;
}
transformed parameters {
  vector[N_obs] mm;
  vector[N_obs] dm;
  vector<upper=0>[N_SN] pt0;
  matrix<lower=0>[N_SN, N_filt] t1;
  matrix<lower=0>[N_SN, N_filt] t2;
  matrix<lower=0>[N_SN, N_filt] td;
  matrix<lower=0>[N_SN, N_filt] tp;
  matrix[N_SN, N_filt] lalpha;
  matrix[N_SN, N_filt] lbeta1;
  matrix[N_SN, N_filt] lbeta2;
  matrix[N_SN, N_filt] lbetadN;
  matrix[N_SN, N_filt] lbetadC;
  matrix[N_SN, N_filt] Mp;
  matrix[N_SN, N_filt] Yb;
  matrix<lower=0>[N_SN, N_filt] V;
  matrix<lower=0>[N_SN, N_filt] M1;
  matrix<lower=0>[N_SN, N_filt] M2;
  matrix<lower=0>[N_SN, N_filt] Md;
  for (l in 1:N_SN) {
    if (duringseason[l] == 1) {
      pt0[l] <- -exp( t0s_h + sig_t0s_h * ( t0s_hSN[l] .* sig_t0s_hSN[l] ));
    } else {
      pt0[l] <- -exp( t0l_h + sig_t0l_h * ( t0l_hSN[l] .* sig_t0l_hSN[l] ));
    }
  }
  for (i in 1:N_filt) {

```



```

for (j in 1:N_SN) {
  t1[j,i] <- exp( log(1) + t_hP[1] + sig_t_hP[1] * (
    t_hF[1,i] * sig_t_hF[1,i]
    + sig_t_hSNF[1,(i-1)*N_SN+j] * t_hSNF[1,(i-1)*N_SN+j]
  ));

  tp[j,i] <- exp( log(10) + t_hP[2] + sig_t_hP[2] * (
    t_hF[2,i] * sig_t_hF[2,i]
    + sig_t_hSNF[2,(i-1)*N_SN+j] * t_hSNF[2,(i-1)*N_SN+j]
  ));

  t2[j,i] <- exp( log(100) + t_hP[3] + sig_t_hP[3] * (
    t_hF[3,i] * sig_t_hF[3,i]
    + sig_t_hSNF[3,(i-1)*N_SN+j] * t_hSNF[3,(i-1)*N_SN+j]
  ));

  td[j,i] <- exp( log(10) + t_hP[4] + sig_t_hP[4] * (
    t_hF[4,i] * sig_t_hF[4,i]
    + sig_t_hSNF[4,(i-1)*N_SN+j] * t_hSNF[4,(i-1)*N_SN+j]
  ));

  lalpha[j,i] <- -1 + ( r_hP[1] + sig_r_hP[1] * (
    r_hF[1,i] * sig_r_hF[1,i]
    + sig_r_hSNF[1,(i-1)*N_SN+j] * r_hSNF[1,(i-1)*N_SN+j]
  ));

  lbeta1[j,i] <- -4 + ( r_hP[2] + sig_r_hP[2] * (
    r_hF[2,i] * sig_r_hF[2,i]
    + sig_r_hSNF[2,(i-1)*N_SN+j] * r_hSNF[2,(i-1)*N_SN+j]
  ));

  lbeta2[j,i] <- -4 + ( r_hP[3] + sig_r_hP[3] * (
    r_hF[3,i] * sig_r_hF[3,i]
    + sig_r_hSNF[3,(i-1)*N_SN+j] * r_hSNF[3,(i-1)*N_SN+j]
  ));

  lbetadN[j,i] <- -3 + ( r_hP[4] + sig_r_hP[4] * (
    r_hF[4,i] * sig_r_hF[4,i]
    + sig_r_hSNF[4,(i-1)*N_SN+j] * r_hSNF[4,(i-1)*N_SN+j]
  ));

  lbetadC[j,i] <- -5 + ( r_hP[5] + sig_r_hP[5] * (
    r_hF[5,i] * sig_r_hF[5,i]
    + sig_r_hSNF[5,(i-1)*N_SN+j] * r_hSNF[5,(i-1)*N_SN+j]
  ));

  Mp[j,i] <- exp(M_h + sig_M_h * (
    M_hF[i] * sig_M_hF[i]
    + sig_M_hSNF[(i-1)*N_SN+j] * M_hSNF[(i-1)*N_SN+j]
  ));

  Yb[j,i] <- Y_h + sig_Y_h * (Y_hSNF[(i-1)*N_SN+j] .* sig_Y_hSNF[(i-1)*N_SN+j]);
  V[j,i] <- V_h * V_hF[i] * V_hSNF[(i-1)*N_SN+j];
}
}
M1 <- Mp ./ exp( exp(lbeta1) .* tp );
M2 <- Mp .* exp( -exp(lbeta2) .* t2 );
Md <- M2 .* exp( -exp(lbetadN) .* td );
for (n in 1:N_obs) {
  real N_SNC;
  int Kc_up;
  int Kc_down;
  real t_exp;
  int j;
  int k;
  real mm_1;
  real mm_2;
  real mm_3;
  real mm_4;
  real mm_5;
  real mm_6;
  j <- J[n];

```

```

    k <- SNid[n];
    t_exp <- ( t[n] - (t0_mean[k] + pt0[k]) ) / (1 + z[k]);
    if (t_exp < 0) {
        mm_1 <- Yb[k,j];
    } else {
        mm_1 <- 0;
    }
    if ((t_exp >= 0) && (t_exp < t1[k,j])) {
        mm_2 <- Yb[k,j] + M1[k,j] * pow(t_exp / t1[k,j] , exp(lalpha[k,j]));
    } else {
        mm_2 <- 0;
    }
    if ((t_exp >= t1[k,j]) && (t_exp < t1[k,j] + tp[k,j])) {
        mm_3 <- Yb[k,j] + M1[k,j] * exp(exp(lbeta1[k,j]) * (t_exp - t1[k,j]));
    } else {
        mm_3 <- 0;
    }
    if ((t_exp >= t1[k,j] + tp[k,j]) && (t_exp < t1[k,j] + tp[k,j] + t2[k,j])) {
        mm_4 <- Yb[k,j] + Mp[k,j] * exp(-exp(lbeta2[k,j]) * (t_exp - t1[k,j] - tp[k,j]));
    } else {
        mm_4 <- 0;
    }
    if ((t_exp >= t1[k,j] + tp[k,j] + t2[k,j]) && (t_exp < t1[k,j] + tp[k,j] + t2[k,j]
        + td[k,j])) {
        mm_5 <- Yb[k,j] + M2[k,j] * exp(-exp(lbetadN[k,j]) * (t_exp - t1[k,j] - tp[k,j]
        - t2[k,j]));
    } else {
        mm_5 <- 0;
    }
    if (t_exp >= t1[k,j] + tp[k,j] + t2[k,j] + td[k,j]) {
        mm_6 <- Yb[k,j] + Md[k,j] * exp(-exp(lbetadC[k,j]) * (t_exp - t1[k,j] - tp[k,j]
        - t2[k,j] - td[k,j]));
    } else {
        mm_6 <- 0;
    }
    dm[n] <- sqrt(pow(dfL[n],2) + pow(V[k,j],2));
    if (t_exp < 0) {
        N_SNc <- 0;
    } else if (t_exp < Kcor_N-2){
        Kc_down <- 0;
        while ((Kc_down+1) < t_exp) {
            Kc_down <- Kc_down + 1;
        }
        Kc_up <- Kc_down+1;
        N_SNc <- Kcor[k,j,Kc_down+1] + (t_exp - floor(t_exp)) * (Kcor[k,j,Kc_up+1]
        - Kcor[k,j,Kc_down+1]);
    } else {
        N_SNc <- Kcor[k,j,Kcor_N];
    }
    mm[n] <- (mm_1+mm_2+mm_3+mm_4+mm_5+mm_6) / (pow(10, N_SNc/(-2.5)));
}
}
model {
    t0s_h ~ normal(0, 0.5);
    sig_t0s_h ~ cauchy(0, 0.1);
    t0l_h ~ normal(log(100), 1);
    sig_t0l_h ~ cauchy(0, 0.1);
    V_h ~ cauchy(0, 0.001);
    Y_h ~ normal(0, 0.1);
    sig_Y_h ~ cauchy(0, 0.01);
    M_h ~ normal(0, 1);
    sig_M_h ~ cauchy(0, 0.1);
    t_hP ~ normal(0,0.1);

```

```

sig_t_hP    cauchy(0, 0.1);
for (i in 1:4) {
  t_hF[i]    normal(prior_t_hF[i], prior_t_hF_s[i]);
  sig_t_hF[i] cauchy(0, 0.1);
  t_hSNF[i]   normal(0,1);
  sig_t_hSNF[i] cauchy(0, 0.1);
}
r_hP    normal(0,1);
sig_r_hP    cauchy(0, 0.1);
for (i in 1:5) {
  r_hF[i]    normal(prior_r_hF[i], prior_r_hF_s[i]);
  sig_r_hF[i] cauchy(0, 0.1);
  r_hSNF[i]   normal(0,1);
  sig_r_hSNF[i] cauchy(0, 0.1);
}
M_hF    normal(0,1);
sig_M_hF    cauchy(0, 0.1);
M_hSNF    normal(0,1);
sig_M_hSNF    cauchy(0, 0.1);
Y_hSNF    normal(0,1);
sig_Y_hSNF    cauchy(0, 0.1);
V_hF    cauchy(0, 0.1);
V_hSNF    cauchy(0, 0.1);
t0s_hSN    normal(0,1);
sig_t0s_hSN    cauchy(0, 0.1);
t0l_hSN    normal(0,1);
sig_t0l_hSN    cauchy(0, 0.1);
fL    normal(mm,dm);
}

```

REFERENCES

- Anderson, J. P., González-Gaitán, S., Hamuy, M., et al. 2014, *ApJ*, **786**, 67
- Bersten, M. C., & Hamuy, M. 2009, *ApJ*, **701**, 200
- Betancourt, M. J. 2013, arXiv:1304.1920
- Betancourt, M. J., & Girolami, M. 2013, arXiv:1312.0906
- Ekström, S., Georgy, C., Eggenberger, P., et al. 2012, *A&A*, **537**, A146
- Foley, R. J., & Mandel, K. 2013, *ApJ*, **778**, 167
- Gelman, A., Carlin, J., Stern, H., et al. 2013, *Bayesian Data Analysis* (3rd ed., Chapman & Hall/CRC Texts in Statistical Science; Boca Raton, FL: Taylor & Francis)
- Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. 2008, *AnApS*, **2**, 1360
- Girolami, M., & Calderhead, B. 2011, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **73**, 123
- Groh, J. H., Meynet, G., & Ekström, S. 2013, *A&A*, **550**, L7
- Heger, A., Fryer, C. L., Woosley, S. E., Langer, N., & Hartmann, D. H. 2003, *ApJ*, **591**, 288
- Hodapp, K. W., Siegmund, W. A., Kaiser, N., et al. 2004, *Proc. SPIE*, **5489**, 667
- Hoffman, M. D., & Gelman, A. 2014, *J. Mach. Learn. Res.*, **15**, 1593
- Jerkstrand, A., Smartt, S. J., Fraser, M., et al. 2014, *MNRAS*, **439**, 3694
- Kaiser, N., Aussel, H., Burke, B. E., et al. 2002, *Proc. SPIE*, **4836**, 154
- Kelly, P. L., & Kirshner, R. P. 2012, *ApJ*, **759**, 107
- Lineweaver, C. H., Fenner, Y., & Gibson, B. K. 2004, *Sci*, **303**, 59
- LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv:0912.0201
- Lunnan, R., Chornock, R., Berger, E., et al. 2014, *ApJ*, **787**, 138
- Magnier, E. 2006, *The Advanced Maui Optical and Space Surveillance Technologies Conference*, ed. S. Ryan (Maui, HI: The Maui Economic Development Board), E50
- Magnier, E. A., Liu, M., Monet, D. G., & Chambers, K. C. 2008, in *IAU Symp. 248, A Giant Step: From Milli- to Micro- Arcsecond Astrometry*, ed. W. J. Jin, I. Platais, & M. A. C. Perryman (Cambridge: Cambridge Univ. Press), 553
- Magnier, E. A., Schlafly, E., Finkbeiner, D., et al. 2013, *ApJS*, **205**, 20
- Mandel, K. S., Foley, R. J., & Kirshner, R. P. 2014, *ApJ*, **797**, 75
- Mandel, K. S., Narayan, G., & Kirshner, R. P. 2011, *ApJ*, **731**, 120
- Mandel, K. S., Wood-Vasey, W. M., Friedman, A. S., & Kirshner, R. P. 2009, *ApJ*, **704**, 629
- Mazzali, P. A., Deng, J., Tominaga, N., et al. 2003, *ApJL*, **599**, L95
- McCrum, M., Smartt, S. J., Rest, A., et al. 2014, arXiv:1402.1631
- Nomoto, K., Kobayashi, C., & Tominaga, N. 2013, *ARA&A*, **51**, 457
- Nomoto, K., Tominaga, N., Umeda, H., Kobayashi, C., & Maeda, K. 2006, *NuPhA*, **777**, 424
- Papaspiliopoulos, O., Roberts, G. O., & Sköld, M. 2007, *StaSc*, **22**, 59
- Rest, A., Stubbs, C., Becker, A. C., et al. 2005, *ApJ*, **634**, 1103
- Sanders, N. E., Soderberg, A. M., Gezari, S., et al. 2014, *ApJ*, **799**, 208
- Sanders, N. E., Soderberg, A. M., Levesque, E. M., et al. 2012a, *ApJ*, **758**, 132
- Sanders, N. E., Soderberg, A. M., Valenti, S., et al. 2012b, *ApJ*, **756**, 184
- Schlaflly, E. F., Finkbeiner, D. P., Jurić, M., et al. 2012, *ApJ*, **756**, 158
- Scolnic, D., Rest, A., Riess, A., et al. 2014, *ApJ*, **795**, 45
- Smartt, S. J. 2009, *ARA&A*, **47**, 63
- Smith, N., Li, W., Filippenko, A. V., & Chornock, R. 2011, *MNRAS*, **412**, 1522
- Stan Development Team. 2014a, *Stan: A C++ Library for Probability and Sampling*, Version 2.2
- Stan Development Team. 2014b, *Stan Modeling Language Users Guide and Reference Manual*, Version 2.2
- Stilp, A. M., Dalcanton, J. J., Warren, S. R., et al. 2013, *ApJ*, **772**, 124
- Stubbs, C. W., Doherty, P., Cramer, C., et al. 2010, *ApJS*, **191**, 376
- Tonry, J., & Onaka, P. 2009, in *Advanced Maui Optical and Space Surveillance Technologies Conference*, ed. S. Ryan (Maui, HI: The Maui Economic Development Board), E50
- Tonry, J. L., Stubbs, C. W., Lykke, K. R., et al. 2012, *ApJ*, **750**, 99
- Urobin, V. P., & Chugai, N. N. 2008, *A&A*, **491**, 507